SUSY Searches with Early LHC data in CMS

Yousi Ma
California Institute of Technology
For the CMS collaboration

14Dec2010

Outline

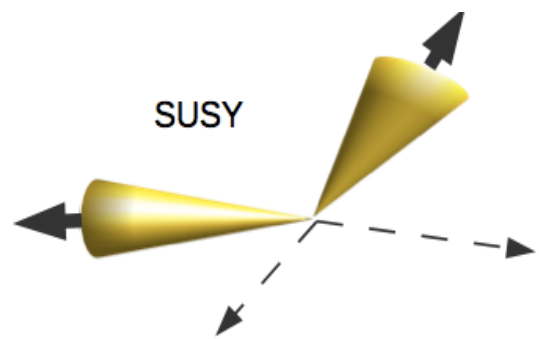
- Data-driven methods in CMS
  - Suppressing QCD with α_T
  - Suppressing QCD with Angular variables
  - Predicting QCD MET with templates
  - Predicting Lepton backgrounds
- Projections for searches
  - All hadronic
  - Like-sign dilepton

Motivations

- SUSY searches involve a broad range of signatures, including jets, leptons, photons, and missing transverse energy (MET)
- Careful control of backgrounds from standard model are needed
- Several methods were developed and tested to:
  - Study QCD backgrounds
  - Control jet-energy mis-measurement
  - Predict background MET distribution
  - Measure contributions from non-prompt leptons/ misidentified hadrons

Suppressing QCD with α_T

- Dimension-less variable $\alpha_T \equiv \frac{p_{T2}}{M_T}$.
- characterizes momentum balance in the event
- allows to suppress multi-jet QCD background
  - based on L. Randall, D. Tucker-Smith, PRL 101 221803 (2008).



$$M_T = \sqrt{2 p_{T1} p_{T2} (1 - \cos\Delta\phi)}.$$

$$\alpha_T = \sqrt{p_{T2}/p_{T1}} / \sqrt{2(1 - \cos\Delta\phi)}.$$

Generalizing to N>2 Jets

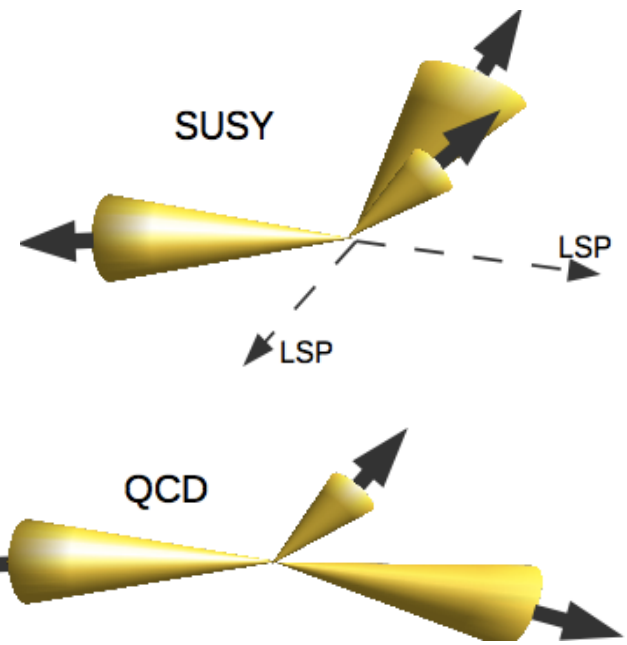
- We generalize the concept of α_T by partitioning a MultiJet system into 2 pseudo-Jets, combining jets vectorally and choosing the unique combination by minimizing ΔH_T

$$H_T = \sum_{\text{jets } j} p_{Tj}$$

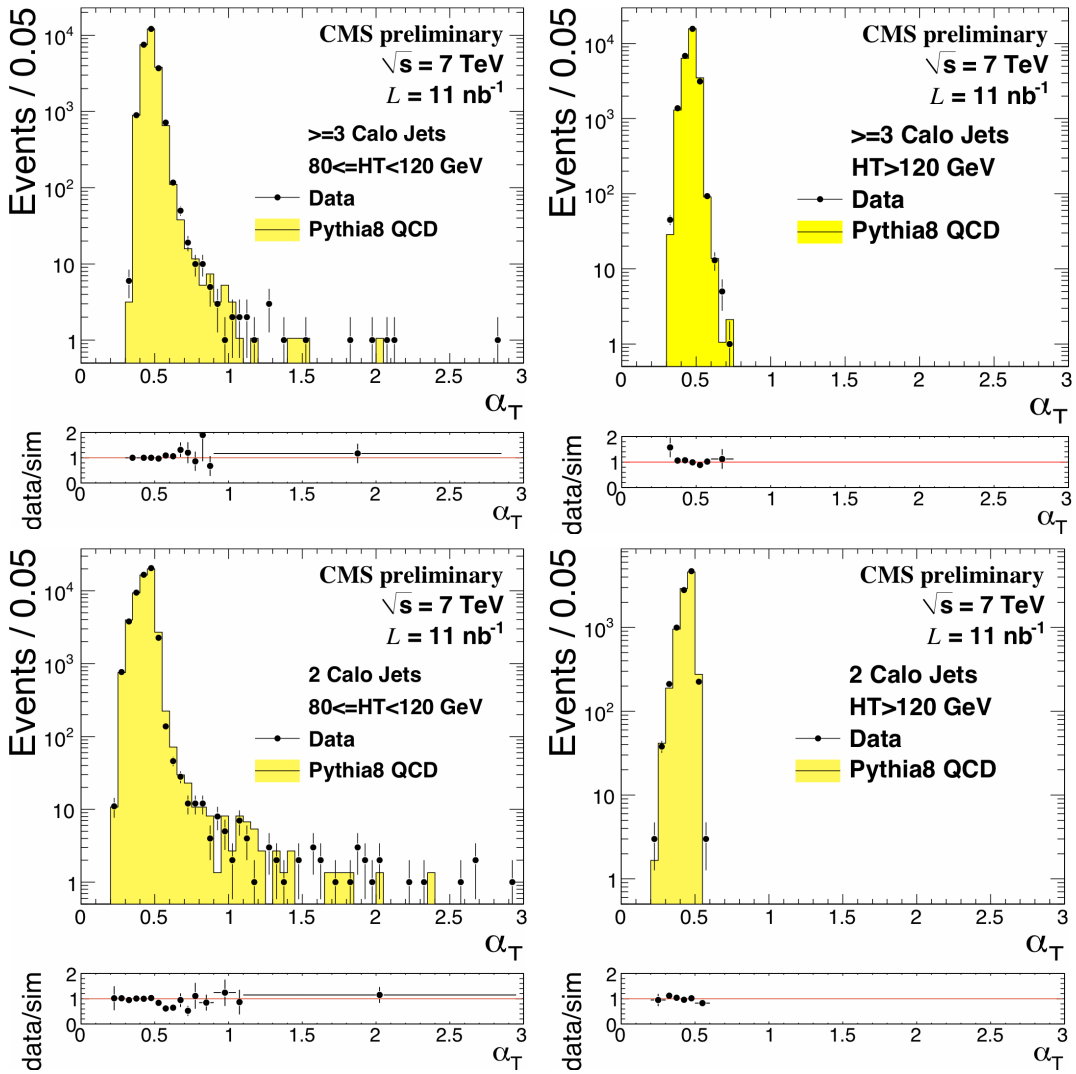
$$\Delta H_T = p_{T\text{pseudojet 1}} - p_{T\text{pseudojet 2}}$$

$$\text{MHT} \equiv |\sum_{\text{jets } j} -\vec{p}_{Tj}|$$

$$\alpha_T = \frac{1}{2}\frac{H_T - \Delta H_T}{\sqrt{H_T^2 - (MHT)^2}}$$



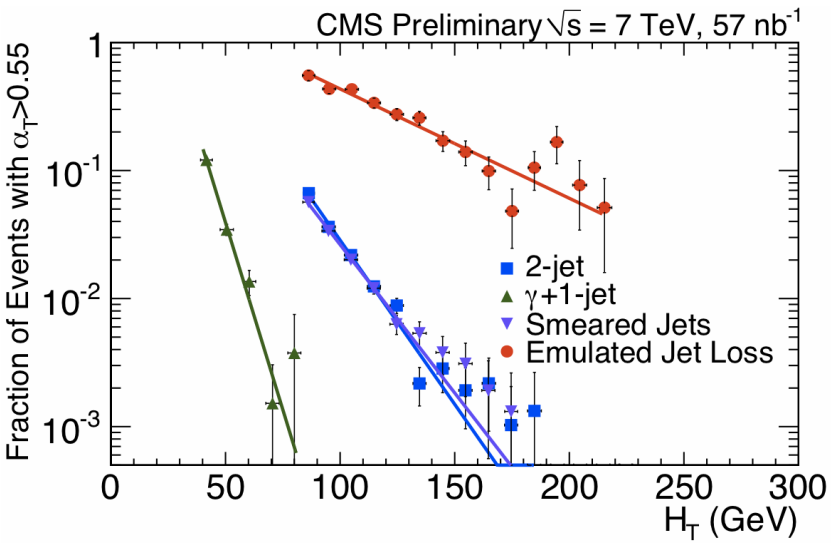
Predicting QCD with α_T

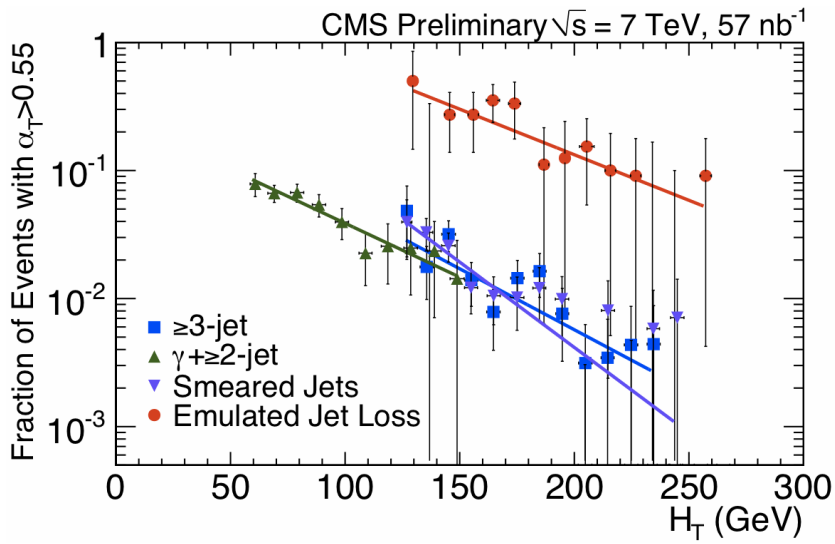


- Trigger: jet 15 GeV (uncor)
- leading jet pT > 40 GeV (corr)
- other jets pT > 20 GeV (corr)
- all jets |η|<3
- jet ID cuts

Predicting QCD with α_T

An exponential decrease of $f(\alpha_T > 0.55)$ in dependence of H_T is observed:

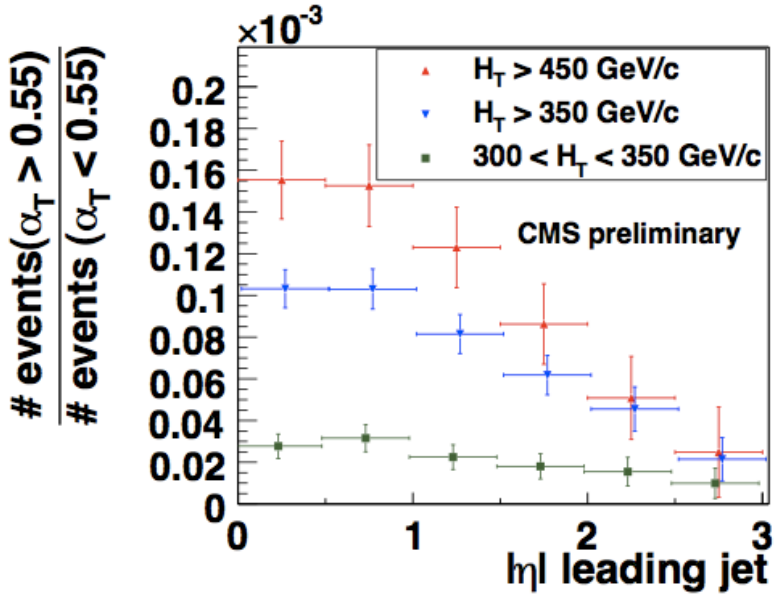


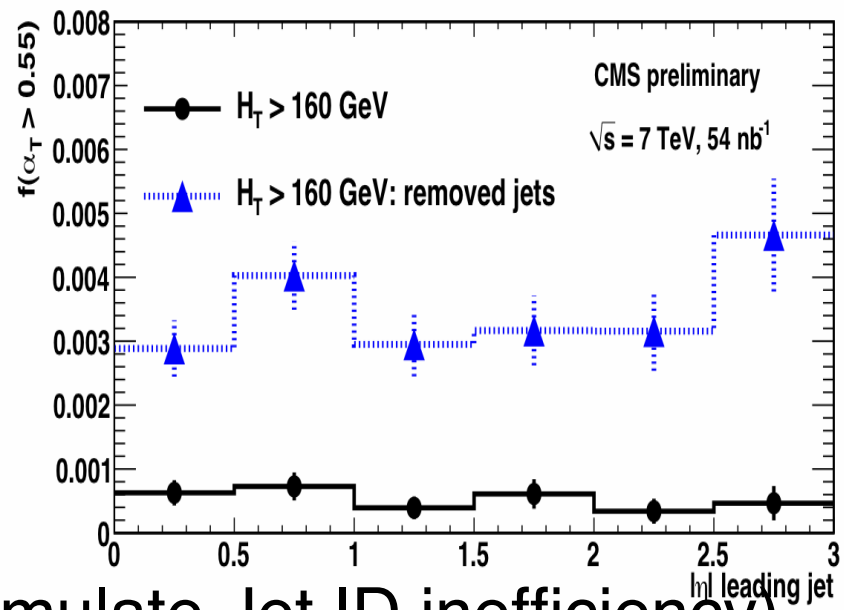


- The jet spectrum is exponential
- The jet-loss probability decreases with p_T
- The effect of a lost jet of fixed p_T decreases with H_T

Monotonic decrease allows us to bound $f(\alpha_T > 0.55)$ at high HT from the values at lower HT

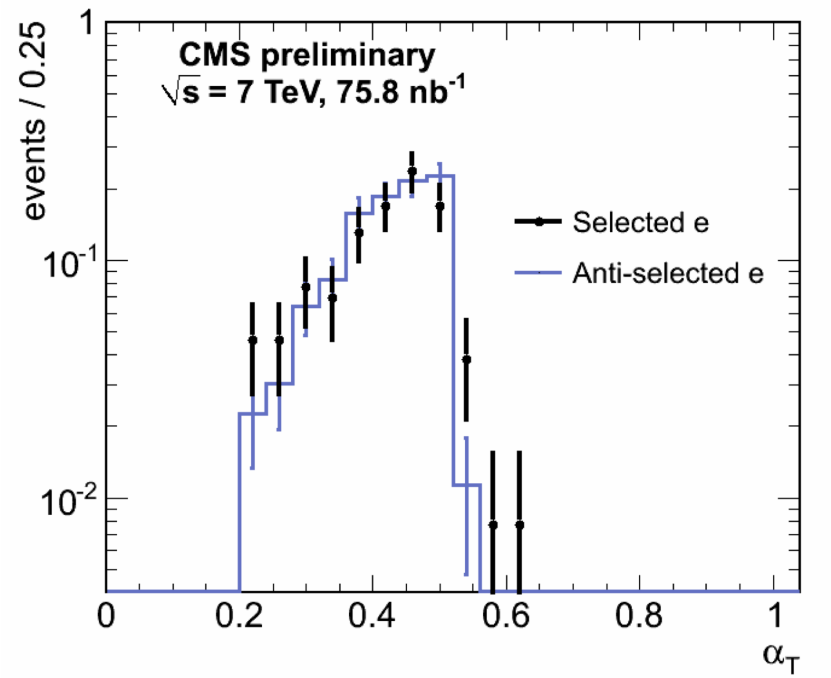
Predicting QCD with α_T

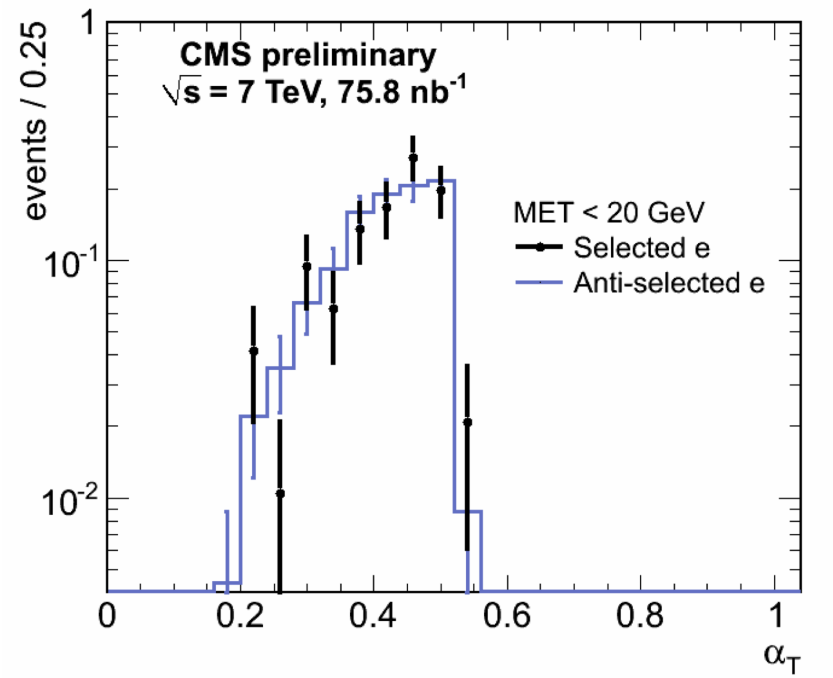




- Even with jet removal (to emulate Jet ID inefficiency) the eta-uniformity of $f(\alpha_T > 0.55)$ is preserved
- Since the $f(\alpha_T > 0.55)$ is also monotonically decreasing, we can use a low HT, and high $|\eta|$ control region to give an upper bound on the background in the central, signal region

Suppressing QCD contribution to l+jets

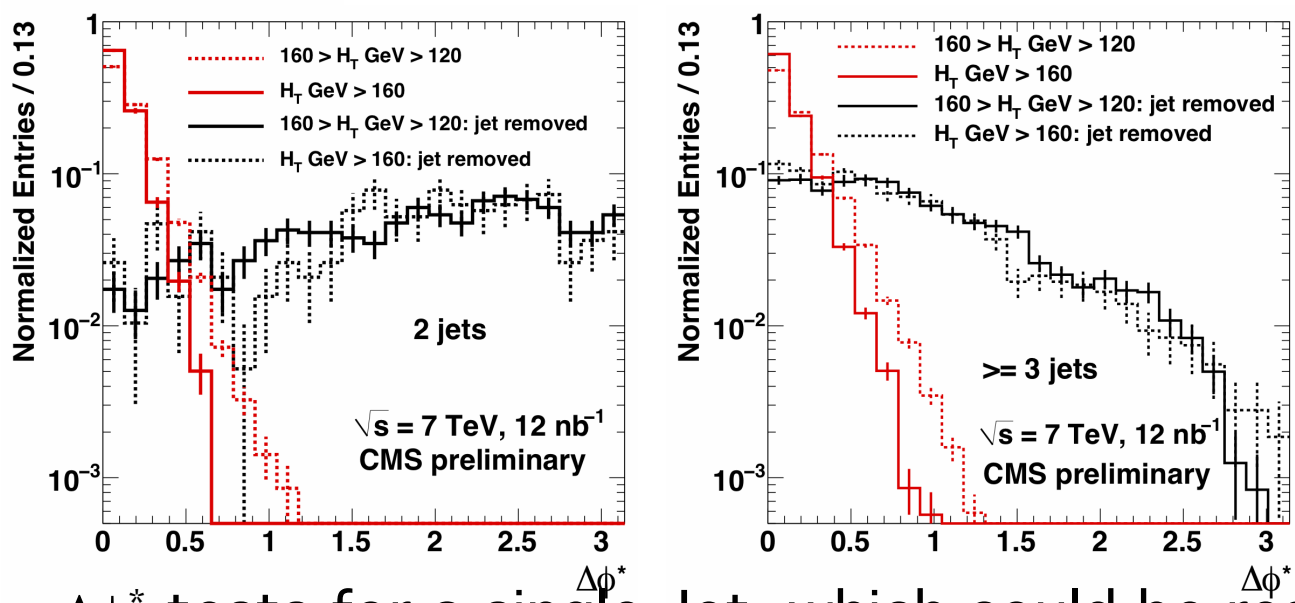




- Form 2 pseudo-jets, from lepton and all jets
- Define α_T in terms of these "jets"
- Anti-selected control sample describes the signal sample, Especially after removing W contamination

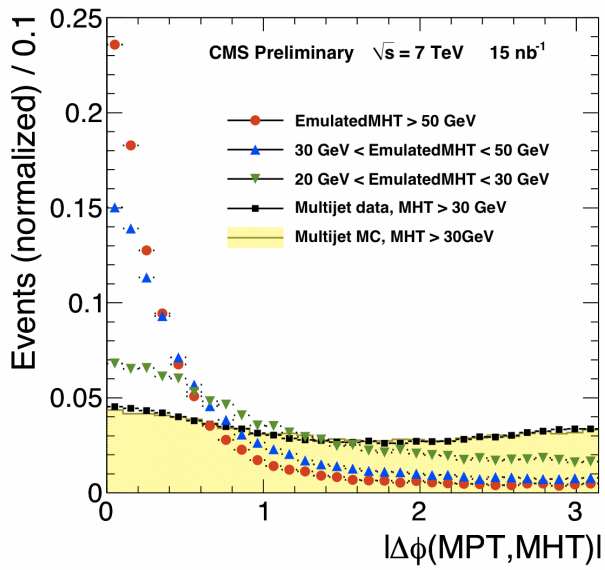
Suppressing QCD with Δϕ(jet, MET)

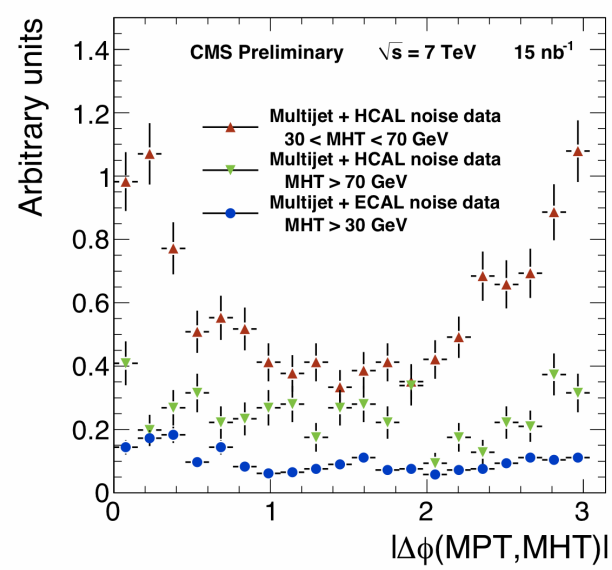
$$\Delta\phi^* \equiv \min_{\text{jets } k}\left(\left|\Delta\phi(\vec{p}_k, -\sum_{\text{jets } i\neq k}\vec{p}_i)\right|\right).$$



- $\Delta\phi^*$ tests for a single Jet, which could be rescaled to balance the event
- Rejects fake MET from Jet Energy mismeasurement

Comparing Track and Calo MET

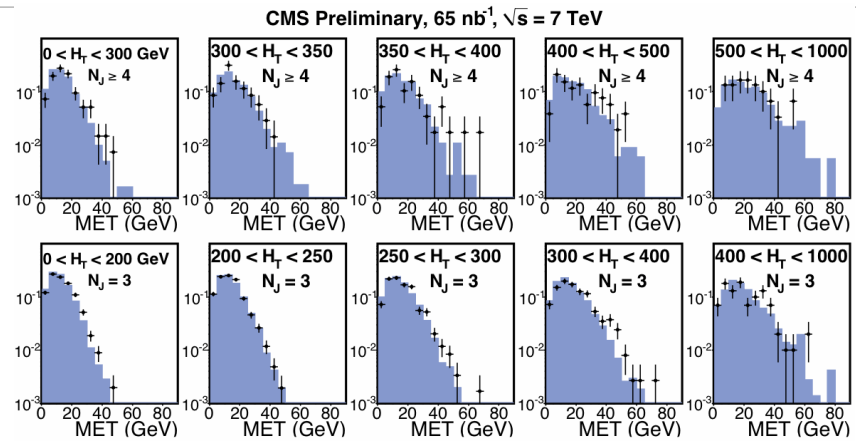


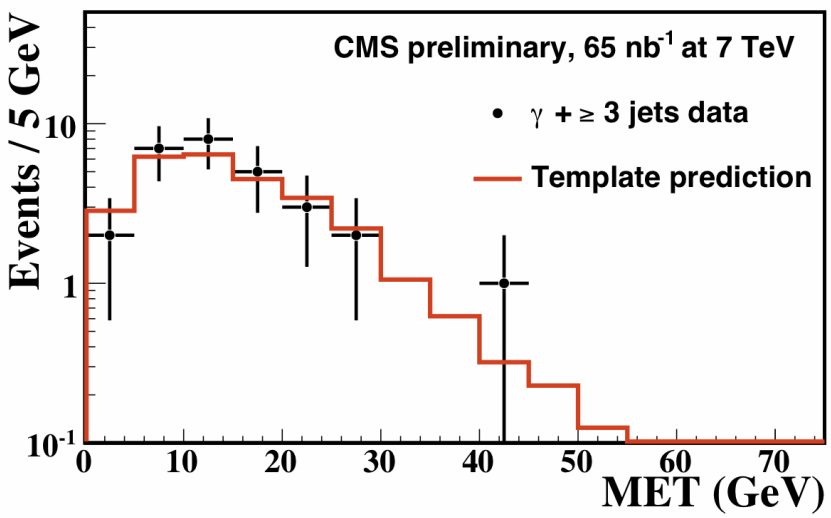


- Since MPT(tracks) and MHT(calo Jet energy) are complementary, expect flat angular distribution for QCD
- For real MET, expect MPT, MHT and MET to be aligned
- Also useful to suppress calorimeter noise, since MPT direction is uncorrelated

Forming MET templates

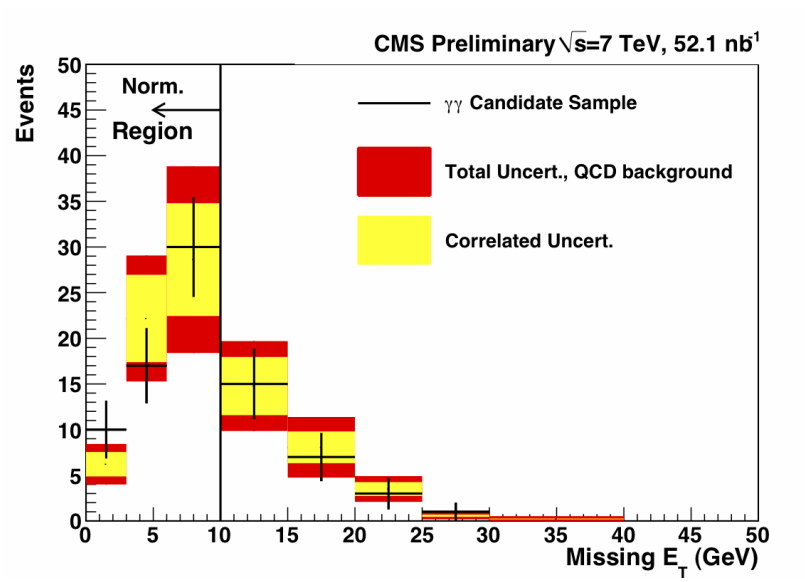
- Form templates for MET using multijet events
- Bin in 2-dimensions:
  - N_{jets}
  - H_T





- Test the MET templates on Photon+jets events
- Good agreement between data and prediction
- N_{data} = 11
- $N_{templates}$ = 12.5

Predicting MET in di-photon fakes



Above 20GeV MET

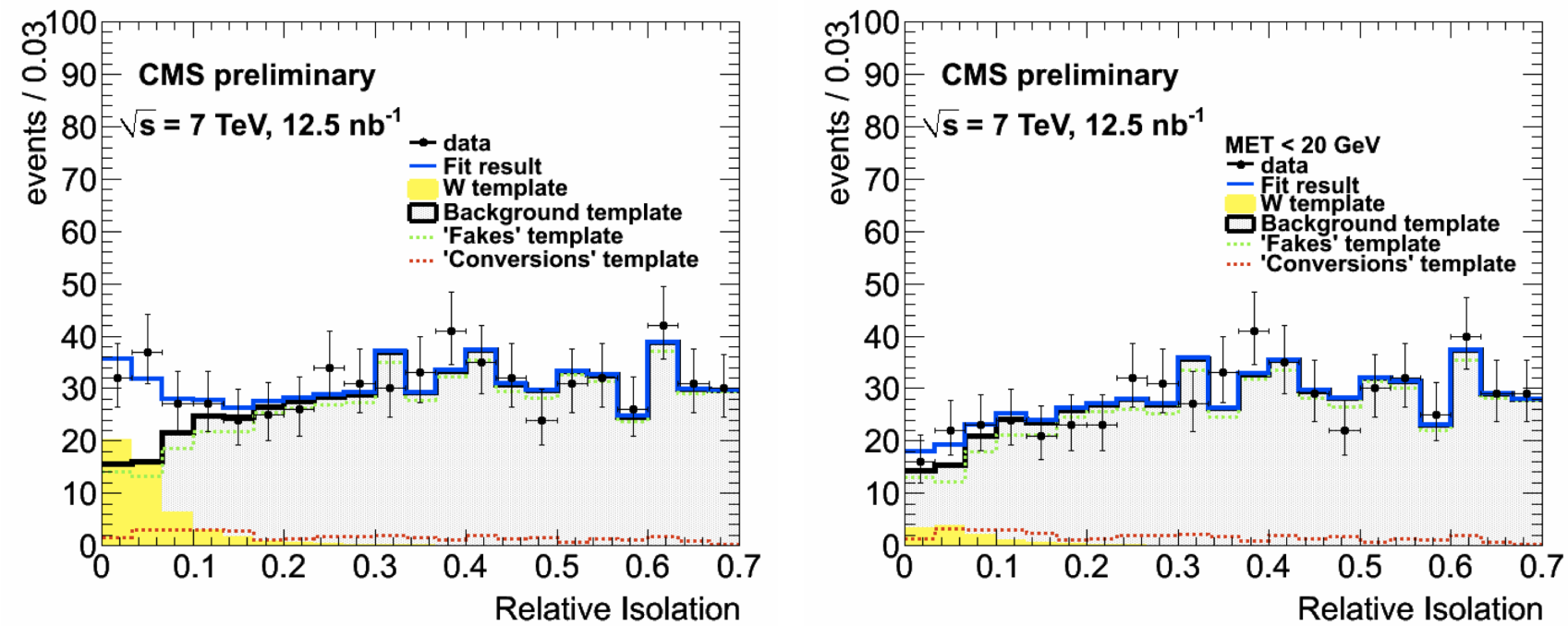
$N_{data} = 4$

$N_{prediction} = 4.2$

- Select a sample with 2 reverse-isolated photon candidates
- Reweight sample such that PT of fake-diphoton system matches that of the signal sample
- Normalize the MET distribution to the yield in the signal region with MET < 10 GeV

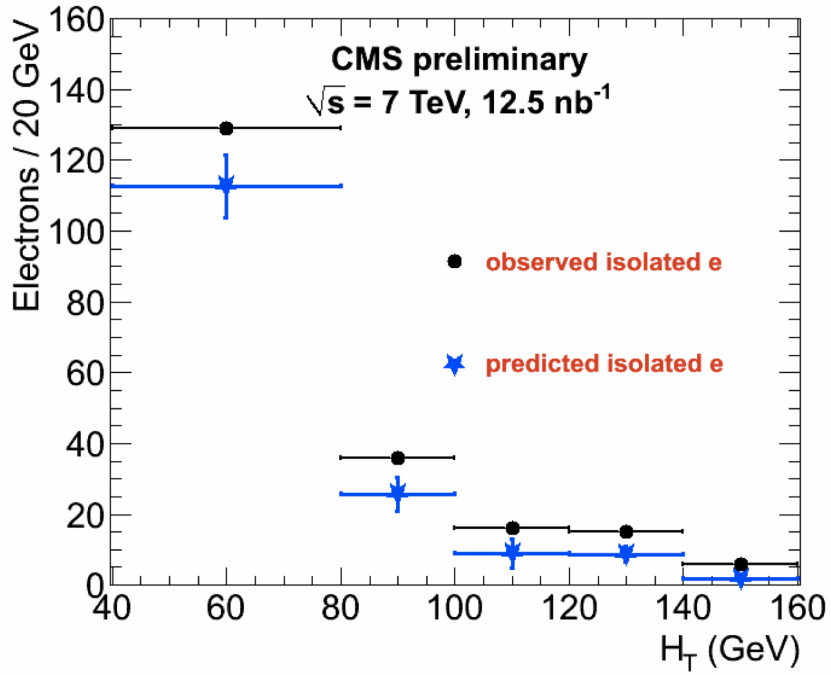
Lepton methods

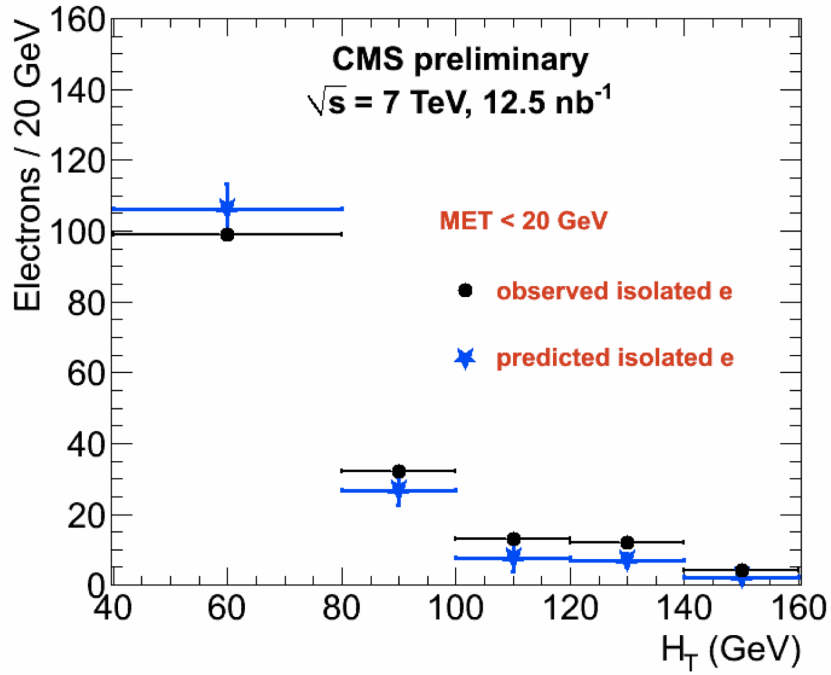
Electron isolation templates



- isolation templates formed from background control samples and lepton MC
- signal sample fit to a linear combination of templates
- W contribution can be suppressed by MET < 20GeV cut

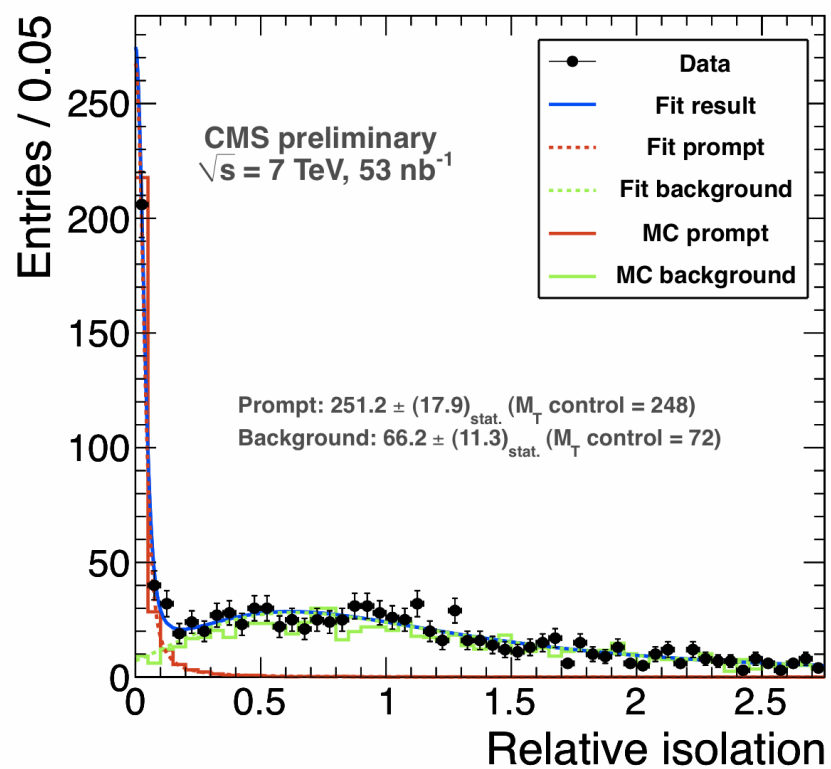
Predicting electron background

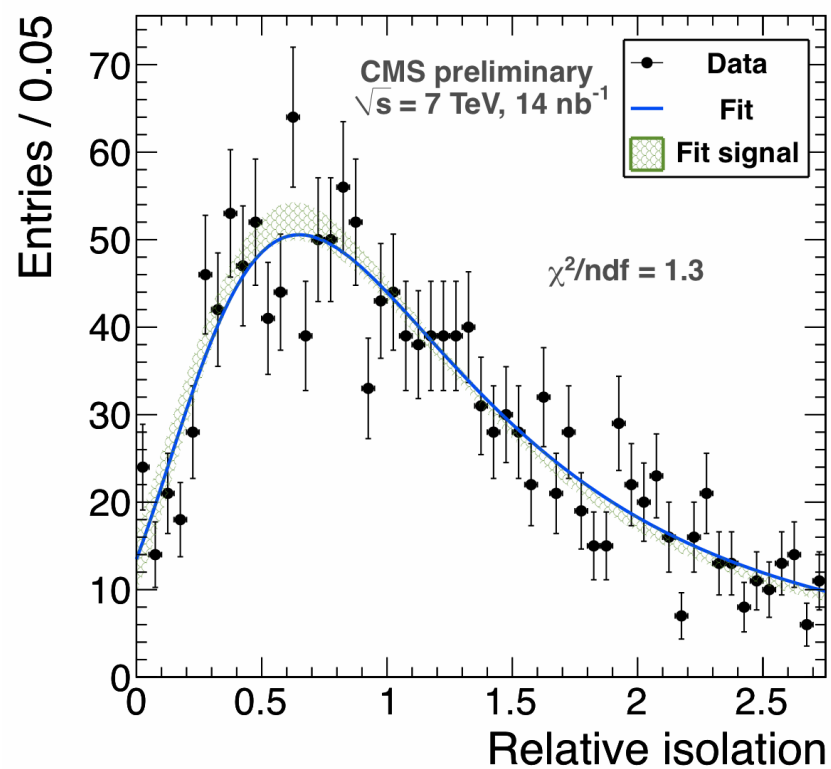




- isolation templates are repeated for different HT bins
- removal of W contribution via MET < 20GeV leads to good agreement between data and MC

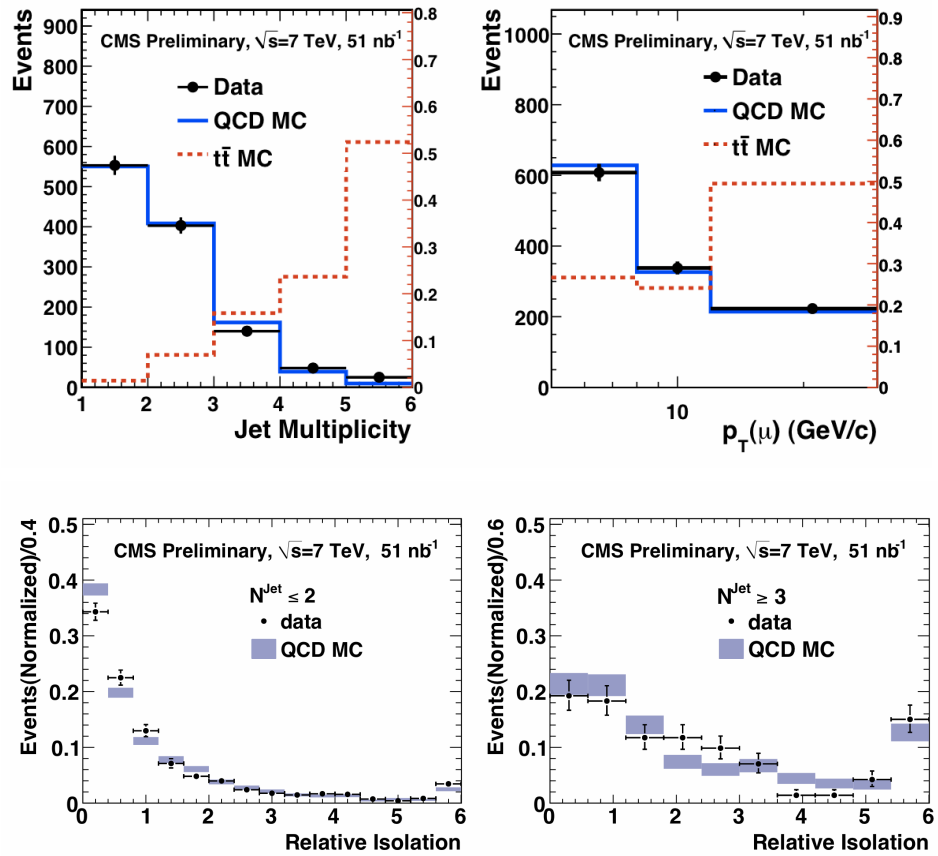
Muon isolation templates





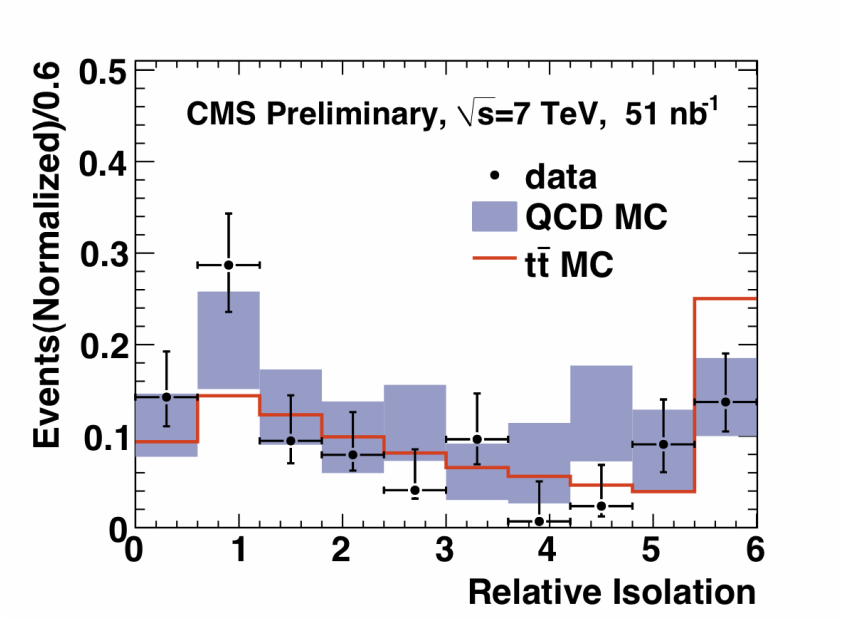
- muon isolation templates (from MC only) also perform well in describing the data
- Again, the non-W component can be studied by required MET < 20GeV

Predicting tt with bb



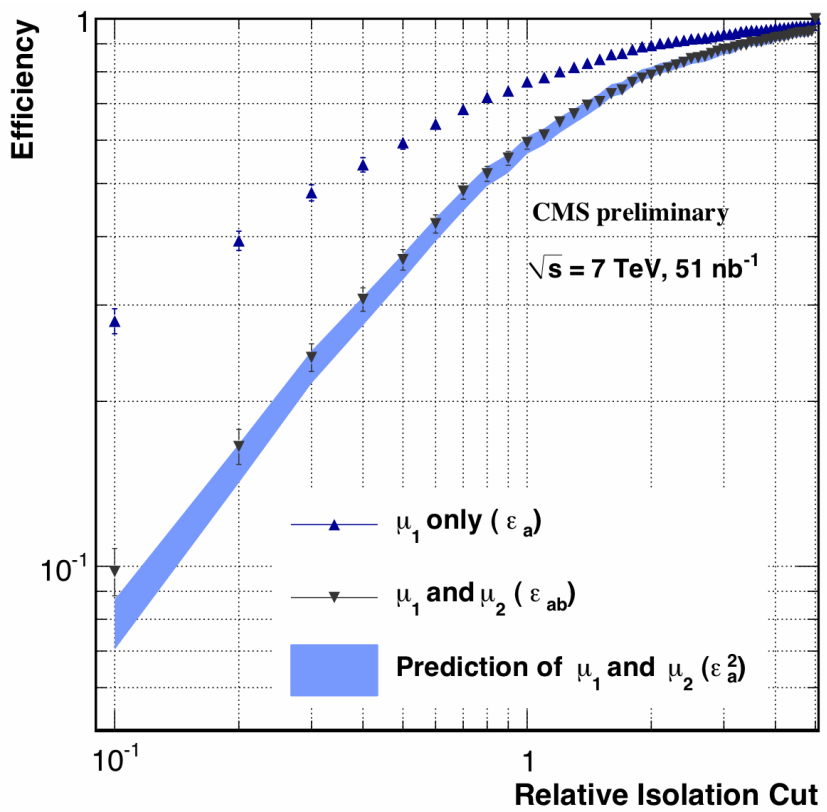
- tt contribution to lepton background is studied by reweighting bb events
- B-tagging identifies opposite side muon from heavy flavor decay
- However, Kinematics of bb and tt are not the same

Reweighted prediction of tt



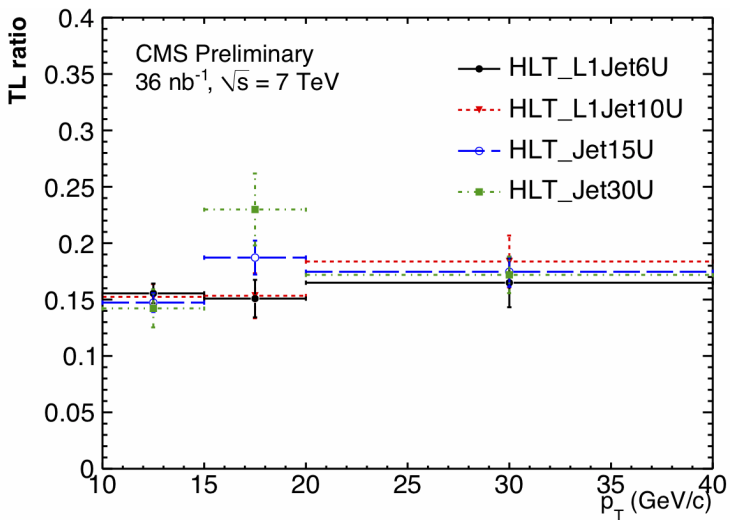
- After reweighting to match event kinematics, QCD and tt MC match within statistics
- Method will allow to use data (reweighted) rather than tt Monte Carlo

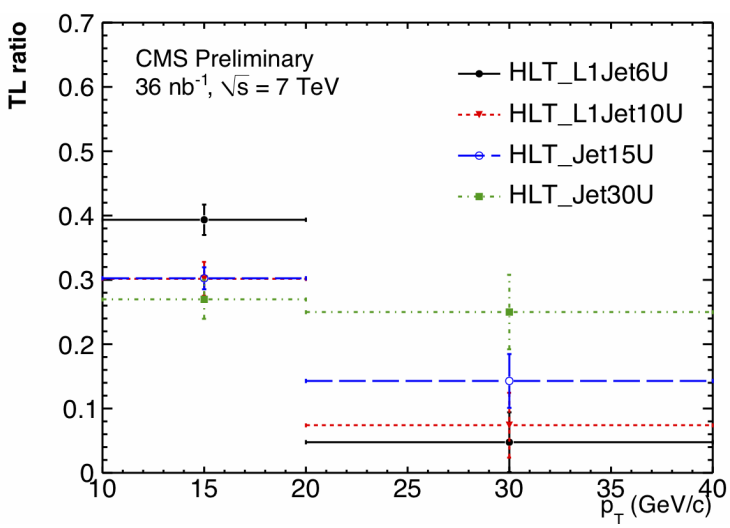
Prediction using factorizability



- Test of factorizability of muon isolation cuts
- Applying all cuts leads to a small sample with possible signal contamination
- Check if individual cuts leads to QCD dominated samples
- Compute the efficiency as a product efficiencies

SS dileptons from tight-loose ratio



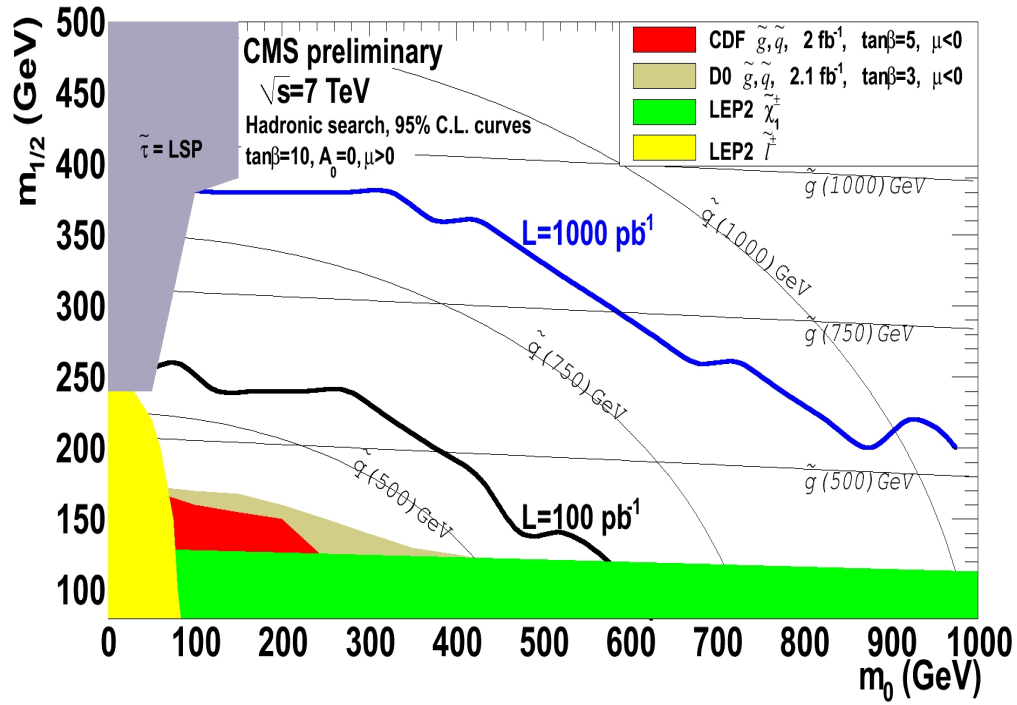


Channel	Predicted	Observed
ee	$0.43^{+0.18}_{-0.14}$	0
$e\mu$	$0.14^{+0.18}_{-0.09}$	1
$\mu\mu$	$0.22^{+0.51}_{-0.18}$	0

- tight identification leaves small statistics for same-sign(SS) dileptons
- find the ratio of SS in tight ID vs. loose ID control samples, as a function of h and pT
- use this ratio on the loose ID signal sample to predict the tight ID SS yield

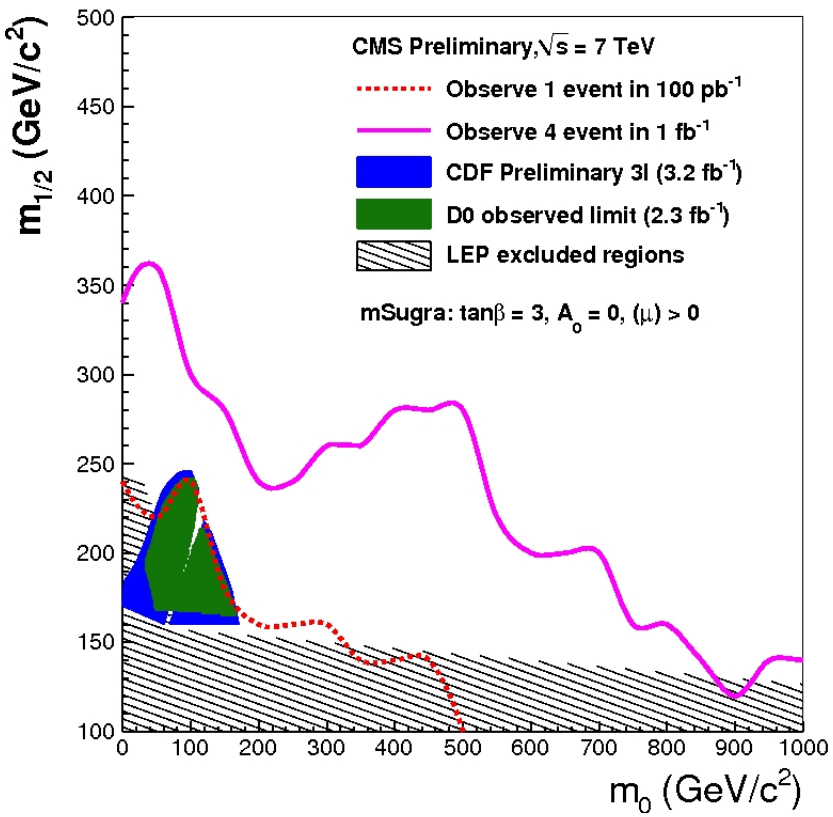
CMS projections

All hadronic search projections



- Veto isolated leptons
- At least 3 jets, $p_T > 50$
- p_T^{MISS} pointing away from leading jets
- $H_T > 400(500)$GeV for 100(1000) pb^{-1}
- $p_T^{MISS} > 225(250)$ for 100(1000)pb^{-1}
- CMS projection only weakly depends on tanβ

Like-sign dileptons projections

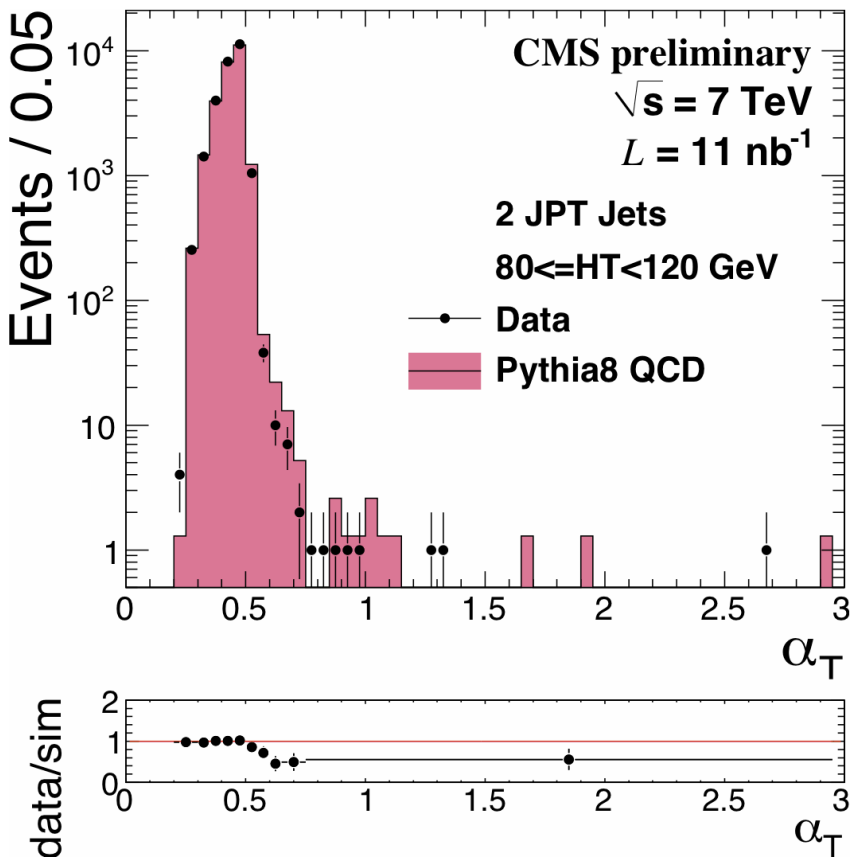


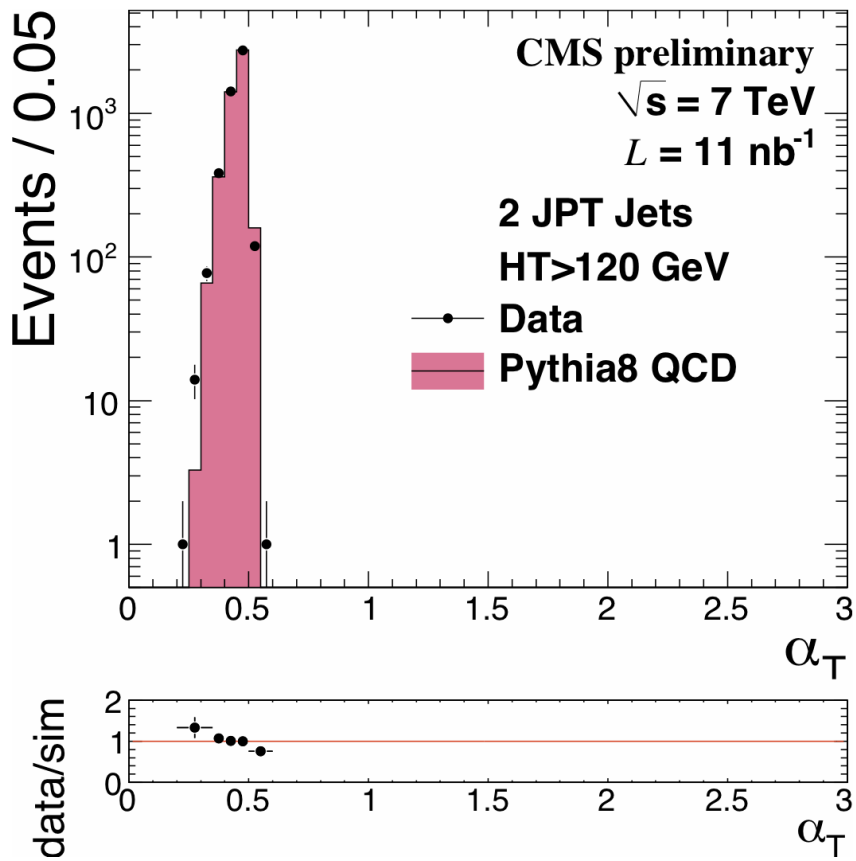
- Two leptons, pT > 20
- At least 3 jets, pT > 30
- HT > 200
- $p_T^{MISS} > 80$
- Very low background expectation for this search

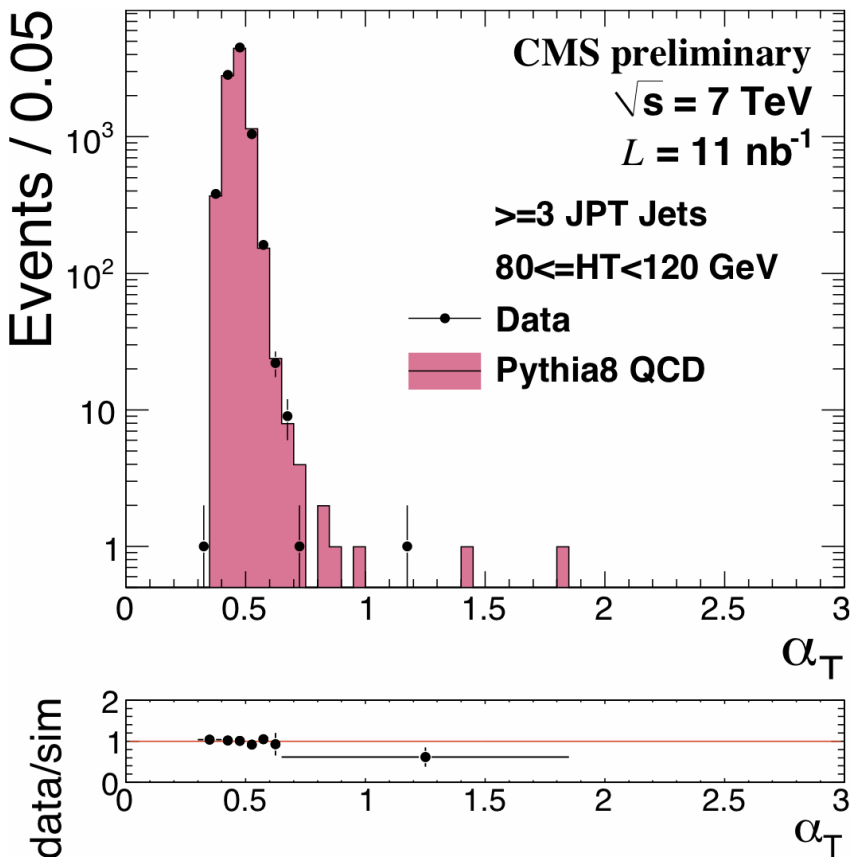
Summary

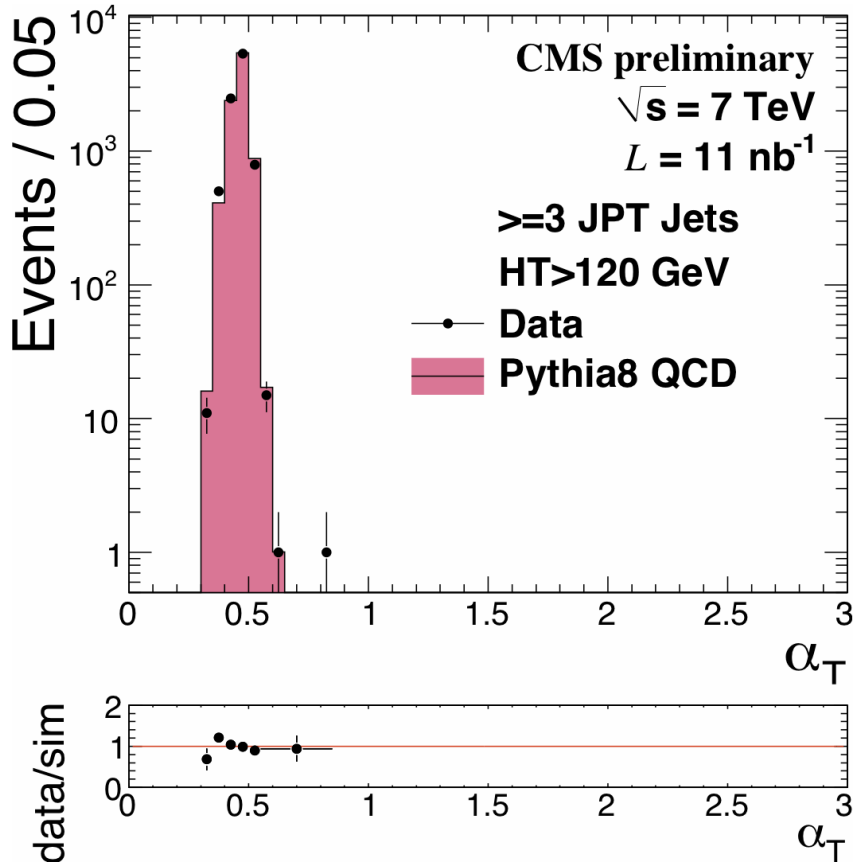
- CMS has studied several data driven methods to suppress and predict QCD backgrounds
- These methods emphasize using one (or more) control samples in data, to estimate the size of an independent sample in signal data
- The statistical precision of these tests will be much improved with the full 2010 dataset, along with a deeper probe of the kinematic regions of interest to full SUSY searches

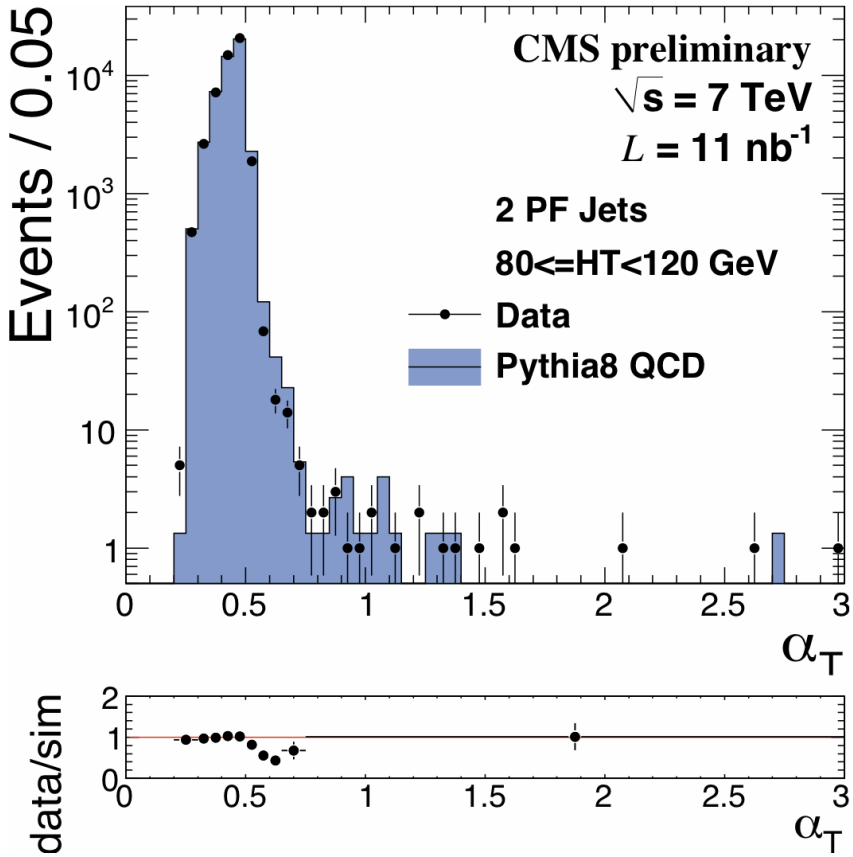
BONUS

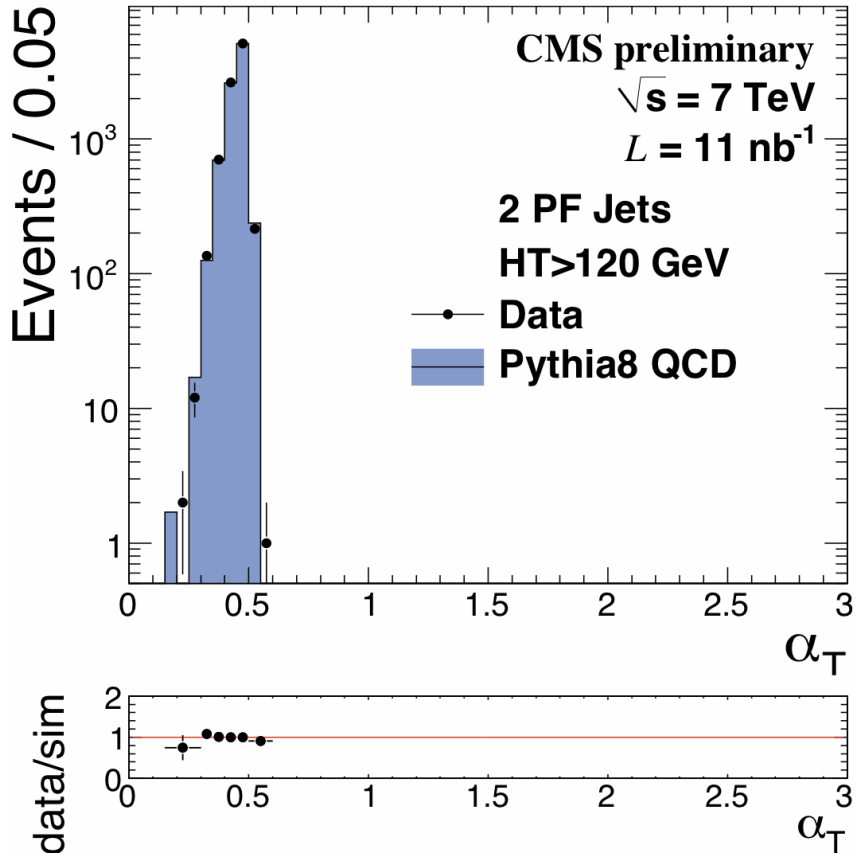

CMS preliminary
√s = 7 TeV
L = 11 nb-1
2 JPT Jets
80<=HT<120 GeV
Data
Pythia8 QCD
Events / 0.05
αT
data/sim

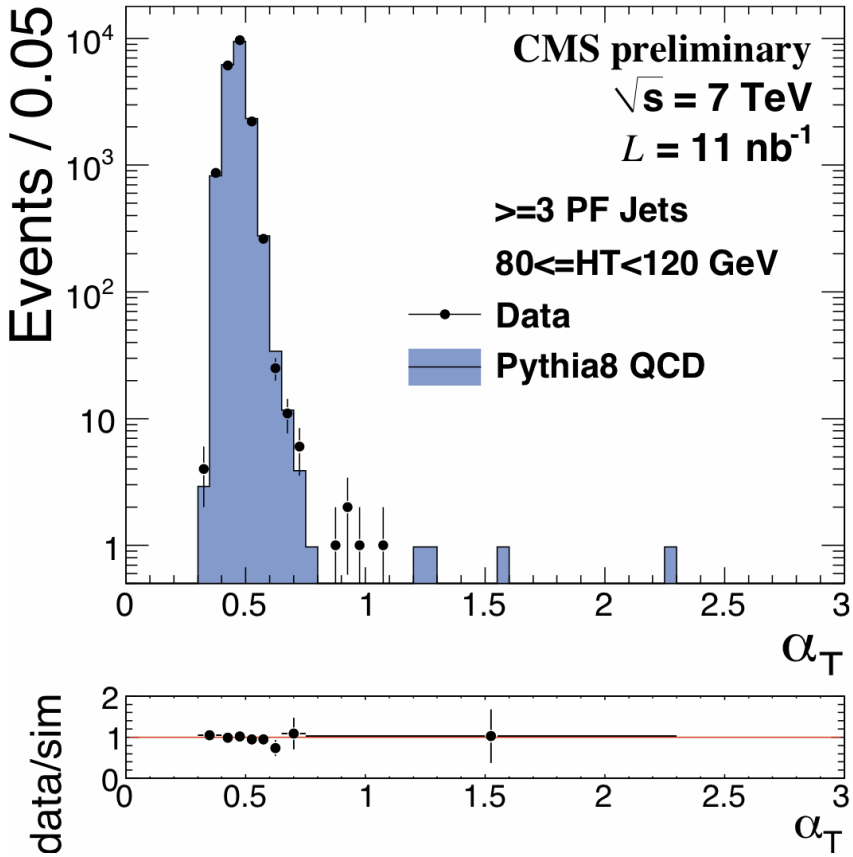

CMS preliminary
√s = 7 TeV
L = 11 nb-1
2 JPT Jets
HT>120 GeV
Data
Pythia8 QCD
Events / 0.05
αT
data/sim

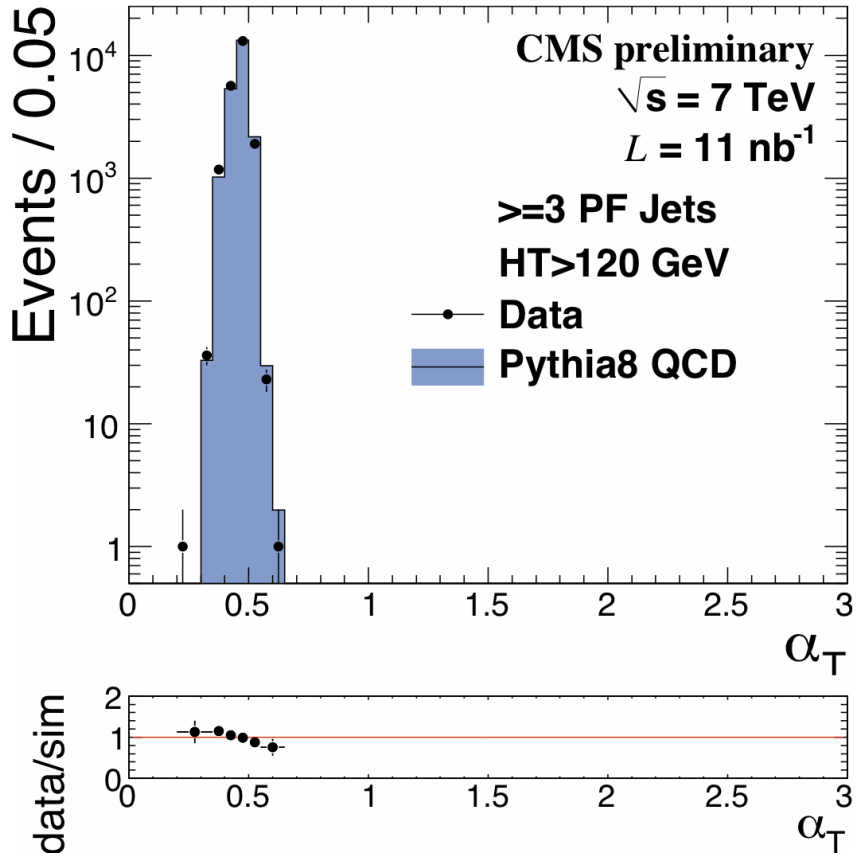

CMS preliminary
√s = 7 TeV
L = 11 nb-1
>=3 JPT Jets
80<=HT<120 GeV
Data
Pythia8 QCD
Events / 0.05
data/sim
αT


CMS preliminary
√s = 7 TeV
L = 11 nb-1
>=3 JPT Jets
HT>120 GeV
Data
Pythia8 QCD
Events / 0.05
data/sim
αT


CMS preliminary
√s = 7 TeV
L = 11 nb-1
2 PF Jets
80<=HT<120 GeV
Data
Pythia8 QCD
Events / 0.05
αT
data/sim


CMS preliminary
√s = 7 TeV
L = 11 nb-1
2 PF Jets
HT>120 GeV
Data
Pythia8 QCD
Events / 0.05
αT
data/sim


CMS preliminary
√s = 7 TeV
L = 11 nb-1
>=3 PF Jets
80<=HT<120 GeV
Data
Pythia8 QCD
Events / 0.05
αT
data/sim


CMS preliminary
√s = 7 TeV
L = 11 nb-1
>=3 PF Jets
HT>120 GeV
Data
Pythia8 QCD
Events / 0.05
αT
data/sim